

OHNLP

Consortium:

- [News](#)
- [Documentation and Downloads](#)
- [Forums](#)
- [Licensing](#)
- [Getting Involved](#)
- [Governance](#)
- [Publications](#)
- [External Links](#)

Follow VKC

Why Natural Language Processing?

The clinical and research medical community creates, manages and uses a wide variety of semi-structured and unstructured textual documents. To perform research, to improve standards of care and to evaluate treatment outcomes easily ? and ideally, in an automated fashion ? access to the content of these documents is required. The knowledge contained in unstructured textual documents (e.g., pathology reports, clinical notes), is critical to achieving all of these goals. For instance, clinical research usually requires the identification of cohorts that follow precisely defined patient- and disease-related inclusion and exclusion parameters. Biomedical NLP systems extract structured information from textual reports, facilitating searching, comparing and summarization.

What is the Goal?

The goal of the Open Health Natural Language Processing Consortium is to establish an open source consortium to promote past and current development efforts and to encourage participation in advancing future efforts. The purpose of this consortium is to facilitate and encourage new annotator and pipeline development, exchange insights and collaborate on novel biomedical natural language processing systems and develop gold-standard corpora for development and testing. The Consortium promotes the open source [UIMA framework and SDK](#) as the basis for biomedical NLP systems. Applications created within UIMA consist of software components (referred to as annotators) and their associated configuration files and external resources. Within the framework, one can also create complete pipelines composed of a sequence of annotators and the data flow between them.